View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Evaluating Asset-Pricing Models Using The Hansen-Jagannathan Bound: A Monte Carlo Investigation

Christopher Otrok
University of Virginia

B. Ravikumar
Pennsylvania State University

Charles H. Whiteman*
University of Iowa

August 2000

Abstract

We conduct Monte Carlo experiments to examine whether the bound proposed by Hansen and Jagannathan (1991) is a useful device for evaluating asset pricing models. Specifically, we use recently developed statistical tests, which are based on a 'distance' between the model and the Hansen-Jagannathan bound, to compute the rejection rates of *true* models. We provide finite-sample critical values for asset pricing models with time separable preferences, and show how they depend upon nuisance parameters—risk aversion and the rate of time preference. Further, we show that the finite-sample distribution of the test statistic associated with the *risk-neutral* case is extreme, in the sense that critical values based on this distribution will deliver type I errors no larger than intended—regardless of risk aversion or the rate of time preference. Extending the analysis to accommodate other preferences, we show that in the state non-separable case, the small-sample distributions of the test statistics are influenced significantly by the degree of intertemporal substitution, but not by attitudes toward risk. For habit formation preferences, the small-sample distributions are strongly influenced by the habit parameter. However, the maximal-size critical values for time-separable preferences are appropriate for habit formation as well as state non-separable preferences. We conclude that with these critical values the HJ bound is indeed a useful evaluation device. We then use the critical values to evaluate three asset pricing models using U.S. data. We find evidence against the time-separable model and mixed evidence on the remaining two models.

* Corresponding author: Charles H. Whiteman, Department of Economics, University of Iowa, Pappajohn Business Building, Iowa city, IA 52242. Emails: cmo3h@virginia.edu, ravikumar@uiowa.edu, whiteman@uiowa.edu. We gratefully acknowledge the support provided for this research by the National Science Foundation under grants SES-0082237 and SES-0882230.

I. Introduction

In asset pricing models, the intertemporal marginal rates of substitution (IMRS) of consumers can be identified given Arrow-Debreu prices from a complete set of securities markets. A typical method of identification is to assume that the IMRS is a parametric function of, say, consumption data, and then test whether the function satisfies the restrictions implied by observed asset price data. An example of this approach is Hansen and Singleton (1982). An alternative approach, advocated by Hansen and Jagannathan (1991), facilitates straightforward study of a broad class of IMRS functions: it utilizes a less stringent restriction that must be met by any proposed IMRS, but, more importantly, this restriction can be calculated independently of the IMRS. The alternative approach does not completely identify the IMRSs, but does provide useful information about them. Specifically, Hansen and Jagannathan (HJ) show that the asset return data imply a lower bound on the volatility of IMRSs of consumers. In doing so, they exploit two simple conditions: the moment condition implied by an asset pricing model (essentially that the expected price of an asset equal the covariance of the asset's payoff with the IMRS), and the requirement (from linear pricing) that the IMRS be a linear function of payoffs. The asset pricing model is said to be consistent with the data if the standard deviation of the IMRS implied by the model is greater than that implied by the HJ bound.[1]

The simple HJ procedure compares just two points and does not account for two types of sampling variability. First, the HJ bound is subject to sampling variability since it is estimated using asset return data. Second, the IMRS is subject to sampling variability since it is estimated using consumption data. Burnside (1994) and Cecchetti, Lam and Mark (1994) develop a statistical test based on the 'distance' between the model IMRS and the HJ bound. (The distance in their case is a normalized difference between the HJ bound and the standard deviation of the IMRS.)[2] Their test reveals whether the asset pricing model is consistent with the data even when the *point* estimate for the model's IMRS volatility does not exceed the *point* estimate for the lower bound.

For the HJ bound to be a useful evaluation device, at the very least, the statistical test should not reject a true model. To be specific, suppose one uses an asset pricing model and observed consumption data to calculate the implied asset returns. Then the test, which is based on the distance between the HJ bound associated with these returns and the volatility of the model IMRS, should not reject the model. A simple method to judge the usefulness of the HJ bound as an evaluation device is to simulate the true

[1] Hansen, Heaton, Luttmer (1995) and Hansen and Jagannathan (1997) assess the extent to which the model is misspecified by calculating pricing errors.

[2] See Cochrane and Hansen (1992) for an alternative distance measure.

model many times and count the number of times the IMRS volatility is below the HJ bound. This is indeed the exercise in Gregory and Smith (1992). Using a model with time-separable isoelastic preferences, Gregory and Smith conclude that the true model is rejected frequently. A more formal statistical evaluation involves calculating rejection rates based on the asymptotic critical values of the Cecchetti-Lam-Mark and Burnside test statistic. This is the exercise in Burnside (1994). He concludes that the true (time-separable) model is *not* rejected frequently. In fact, the statistical test, according to Burnside, is especially good if one uses finite-sample critical values instead of asymptotic critical values.

Our paper adds to the existing literature along a number of dimensions. First, we show how the finite-sample critical values of the test statistic depend upon nuisance parameters—risk aversion and the rate of time preference—in the time-separable asset pricing model. We show that the finite-sample distribution of the test statistic associated with the *risk-neutral* case is extreme, in the sense that critical values based on this distribution will deliver type I errors no larger than intended—regardless of risk aversion or rate of time preference. Second, we provide finite-sample critical values appropriate for extending the HJ assessment procedure to environments with time non-separable and state non-separable preferences. In the state non-separable case, the small-sample distributions of the distance test statistics are influenced significantly by the degree of intertemporal substitution, but not by attitudes toward risk. For time non-separable preferences, the small-sample distributions are strongly influenced by the habit formation parameter. However, we show that the maximal-size critical values for time-separable preferences are appropriate for habit formation as well as state non-separable preferences; i.e., we provide one set of critical values that is applicable to all three forms of preferences and all parameter values. With these critical values, we conclude that the HJ bound is indeed a useful evaluation device. Third, we use our small-sample critical values to test the time-separable model and the state and time non-separable models using quarterly U.S. data from 1947-1997. We find evidence against the time separable model, and mixed evidence with respect to the state and time non-separable models.

The next three sections develop necessary background for implementing the HJ bound. We review the derivation of the bound in Section II. Then in Section III, we follow Burnside (1994) and show how to account for sampling variability in tests based on the bound. In Section IV, we describe model economies with three different forms for preferences—time separable, state non-separable, and time non-separable. Readers familiar with the asset pricing literature may wish to skip Sections II-IV, though the notation for the rest of the paper is set in these sections.

We describe our Monte Carlo procedure for assessing the HJ evaluation device for the time-separable case in Section V. We show how the finite-sample distributions of the test statistics depend upon risk aversion. We extend these calculations to the non-separable cases in Section VI. In Section VII, we use a high-dimension Markov process for consumption to provide very accurate critical values from this extreme distribution; these critical values will be useful in applied work. In Section VIII we test the three asset pricing models using actual data and the critical values from Section VII.

II. The Hansen-Jagannathan Bound

In this section, we review the derivation of the volatility bound presented in Hansen and Jagannathan (1991) for two cases. In the first case, there are n-1 risky assets and one riskless asset; in the second, there are n-1 risky assets and *no* riskless asset. Our derivation of the bound differs somewhat from the original, in ways that will help us contrast the two cases.

II.1 One Riskless Asset and n-1 risky assets

Let p^f denote the current price of an asset that pays one sure unit of consumption in the next period, and let q denote the $(n-1) \times 1$ price vector of risky assets with payoffs x in the next period. Consider two intertemporal marginal rates of substitution ("pricing kernels" or "stochastic discount factors") m and m* that price the assets according to

(1) $EQ = EXm = EXm^*$

where $Q = [p^f\ q']'$, and $X = [1\ x']'$. Equation (1) is the unconditional version of the standard asset pricing model Euler condition equating the expected marginal cost (EQ) and marginal benefit (EXm) of delaying consumption one period. Then

(2) $EX(m-m^*) = 0.$

Since the first element of X is the unit payoff, from equation (1) we know that

(3) $Em = Em^* = Ep^f \equiv v.$

Let P be the linear space given by $\{X'c:\ c \text{ in } \Re^n\}$, and suppose that m* is in P, thereby ensuring linear pricing for m*. (For example, an m* in P satisfying (1) is $m^* = X'c$, where $c = [EXX']^{-1}EQ$.) Define the "error" ε such that

(4) $m = m^* + \varepsilon.$

Note that from (2), $EX\varepsilon = 0$. Thus m* is the linear least squares projection of m onto P. Then

$$\text{var}(m) = \text{var}(m^*) + \text{var}(\varepsilon) + 2\text{cov}(m^*,\varepsilon).$$

Consider the covariance term:

$$\text{Cov}(m^*,\varepsilon) \quad = \quad Em^*\varepsilon - Em^*E\varepsilon$$

$$= \quad Em^*\varepsilon \quad \text{(since } Em = Em^* \text{ implies } E\varepsilon = 0).$$

Since m* is in X and ε is orthogonal to X, we must have $Em^*\varepsilon = 0$, and hence, $Cov(m^*,\varepsilon) = 0$. Thus, we have

$$var(m) = var(m^*) + var(\varepsilon) \geq var(m^*),$$

meaning that the lower bound on the variance of m is that of m*.

To find this lower bound, note that since m* is a linear combination of the elements of X, it can be written in the form

(5) $$m^* = v + (x - Ex)'\beta$$

for β in $\Re^{n-1}$. To obtain the vector β, subtract v from both sides, multiply by (x-Ex) and take expectations:

(6) $$E(x - Ex)(m^* - v) = E(x - Ex)(x - Ex)'\beta = \Sigma\beta$$

where $\Sigma = E(x - Ex)(x - Ex)'$ is the covariance matrix of payoffs on the risky assets. A more convenient expression for the left-hand-side of (6) can be derived as follows. First, from (1), $EXm^* = EQ$, so we have

$$EXm^* - vEX = EQ - vEX$$

which implies

$$EX(m^* - v) = EQ - vEX .$$

Now subtract zero [$= EXE(m^* - v)$] from the left-hand-side to obtain

$$EX(m^* - v) - EXE(m^* - v) = EQ - vEX$$

or

(7) $$E(X - EX)(m^* - v) = EQ - vEX = \begin{bmatrix} Ep^f \\ Eq \end{bmatrix} - v\begin{bmatrix} 1 \\ Ex \end{bmatrix}.$$

Since the first element of X is 1, equation (7) implies

$$E(x - Ex)(m^* - v) = Eq - vEx.$$

Therefore, we can solve (6) for β as

(8) $$\beta = \Sigma^{-1}(Eq - vEx).$$

Thus,

$$\begin{aligned} var(m^*) &= E(m^* - v)'(m^* - v) \\ &= \beta' E(x - Ex)(m^* - v) && \text{(from equation (5))} \\ &= \beta'\Sigma\beta && \text{(from equation (6)).} \end{aligned}$$

Substituting for β from (8), we can write

(9) $$var(m^*) = (Eq - vEx)'\Sigma^{-1}(Eq - vEx).$$

Note that the lower bound is just a point in $\Re_+$.

To implement the bound, it is common to rewrite the right-hand-side of (9) in terms of gross returns. To accomplish this we can normalize q to be a vector of ones and x to be the vector of gross returns. That is, we define each risky asset as carrying a unit price for the stochastic payoffs represented by gross returns. The bound is then written as:

(10) $\quad \text{var}(m^*) = (\iota - vER)'\Omega^{-1}(\iota - vER)$

where R is now the vector of gross returns on the n-1 risky assets, ι is an n-1 vector of ones, and Ω is the covariance matrix of risky-asset returns.

II.2 No riskless asset and n-1 risky assets

Equation (1) now needs to be modified to

(1') $\quad Eq = Exm.$

Let M be a linear space given by $\{x'c: c \text{ in } \Re^{n-1}\}$. Since M does not have a riskless unit payoff, augment M with linear combinations of x and the unit payoff so that we have the linear space P. Pick an arbitrary v $\in \Re$ to denote the (unknown) price of the unit payoff. We seek an m_v with mean v that also satisfies (1'). For all m's such that Em = Em_v we can write (1') as

$$\begin{bmatrix} v \\ Eq \end{bmatrix} = EXm = EXm_v .$$

This system is analogous to (1), and versions of equations (2)-(9) follow immediately with m_v in place of m*. Unlike the case with a riskless asset, the lower bound is not just a point in $\Re_+$ since v is not pinned down by the price of the riskless asset; var(m_v), is a function of the arbitrarily picked v. Thus, by picking different v's we generate a lower bound *frontier*.

As above, implementation of the HJ bounds typically uses the return version of the bound frontier, which is

(10') $\quad \text{var}(m_v) = (\iota - vER)'\Omega^{-1}(\iota - vER)$.

III. Distance to the Bound

The HJ bound may be used to assess the empirical success of a model as follows. Data on returns are used to estimate the mean return vector and covariance matrix appearing in (10) or (10'). Then the IMRS implied by the model is calculated using data on consumption, leisure, etc. and parametric assumptions regarding discount factors, risk aversion, etc. (Examples are presented in the next section.)

Then the variance of the sample IMRS is compared to the bound (10 or 10'), and the model is rejected if the volatility of the IMRS is less than the bound.

This model assessment device requires estimation of the mean and standard deviation of the IMRS, and the mean return vector and the variance-covariance matrix of returns. These estimates are of course subject to sampling variability. To account for this, Cochrane and Hansen (1992), Burnside (1994), Cecchetti, Lam, and Mark (1994), and Hansen, Heaton and Luttmer (1995) consider the sampling distribution of the vertical distance to the bound using a GMM estimator. (The vertical distance refers to the difference between the HJ bound and the standard deviation of the model IMRS.) This allows sampling variability to be taken into account in determining what it means to be "close enough" to the bound. The derivation of the relevant statistics presented below borrows heavily from Burnside (1994) and is included here mainly for completeness. We again focus on two cases: the first has a riskless asset and only one risky asset (instead of n-1); the second has only a single risky asset.

III.1 One Riskless Asset and One Risky Asset

Consider estimating the five moments of interest by simple calculation of the associated sample moments. (There is virtually no "generalization" in this implementation of the "method of moments".) The moment conditions are

(11) $E[p_t^f - \alpha_0] = 0$

(12) $E[m_t - \alpha_1] = 0$

(13) $E[(m_t - \alpha_1)^2 - \alpha_2] = 0$

(14) $E[R_t^e - \alpha_3] = 0$

(15) $E[(R_t^e - \alpha_3)^2 - \alpha_4] = 0$.

Since there are exactly as many moments as parameters, the parameters are exactly identified. Let $f(y_t, \alpha) = 0$ represent the above moment conditions, where y_t is the T×3 matrix with data on gross equity returns, price of the riskless asset, and the representative agent's IMRS, and α is the 5x1 vector of moments.[3] The GMM estimates are derived by choosing α to minimize a quadratic form in the sample

[3] In the case with the riskless asset, recall that the mean IMRS is the expected price of the riskless asset. We do not explicitly impose the restriction $Ep^f = Em$. There are no degrees of freedom available to accomplish this, since once preferences are specified, the sample $\{m_t\}$ process (and thus the mean) is completely determined by the consumption data.

average of $f(y_t, \alpha)$. For details, see Hansen (1982). We denote the estimator of α by $\hat{\alpha}$, and the asymptotic covariance matrix of the estimator of α by $V(\hat{\alpha})$.[4]

In the case of one riskless and one risky asset, the bound in (10) reduces to:

$$\text{(16)} \qquad \text{var}(m^*) = \frac{\left(1 - Ep^f ER^e\right)^2}{\text{var}(R^e)}$$

which implies

$$\text{std}(m^*) = \sqrt{\frac{\left(1 - Ep^f ER^e\right)^2}{\text{var}(R^e)}} \,.$$

A Wald statistic based on the vertical distance to the HJ bound in mean-standard deviation (of IMRS) space is used to examine the HJ bound as an evaluation device. The distance to the bound is given by:

$$\text{(17)} \qquad \gamma_1 \equiv h(\alpha) = \sqrt{\alpha_2} - \sqrt{\frac{(1 - \alpha_0 \alpha_3)^2}{\alpha_4}}$$

where α_0 is the mean price of the riskless asset, α_2 is the variance of the agent's IMRS, α_3 is the mean return on equity, and α_4 is the variance of the equity return.

Under the regularity conditions in Hansen (1982) the asymptotic variance of γ_1 can be estimated with:

$$\text{(18)} \qquad \hat{\sigma}^2_{\gamma_1} = \frac{\partial h(\hat{\alpha})}{\partial \alpha} V(\hat{\alpha}) \frac{\partial h(\hat{\alpha})'}{\partial \alpha} \,.$$

Under the null that the population distance to the bound is zero, we have the following Wald statistic:

$$\text{(19)} \qquad Z_1 \equiv \sqrt{T}\left(\frac{\hat{\gamma}_1}{\hat{\sigma}_{\gamma_1}}\right) \xrightarrow{d} N(0,1) \,.$$

The usefulness of HJ bound as an evaluation device depends on the finite-sample properties of this Wald statistic.

III.2 No Riskless Asset and One Risky Asset

---

[4] In particular, $V(\alpha) = (D_T' S_T^{-1} D_T)^{-1}$, where

$$S_T = \frac{1}{T}\sum_{t=1}^{T} f(y_t,\alpha) f(y_t,\alpha)' + \sum_{i=1}^{n}\left[1 - \frac{i}{n+1}\right] \times \left[\frac{1}{T}\sum_{t=1+i}^{T} f(y_t,\alpha) f(y_{t-i},\alpha)' + \frac{1}{T}\sum_{t=1+i}^{T} f(y_t,\alpha) f(y_{t+i},\alpha)'\right],$$

$D_T = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial f(y_t,\alpha)}{\partial \alpha}$, and n is the number of lags in the Newey-West covariance estimator. In our estimation we use n = 6.

When there is no riskless asset, the derivation of the Wald statistic is similar, but uses an alternative set of moment conditions. The moment conditions are:

(20) $E[m_t - \alpha_1] = 0$

(21) $E[(m_t - \alpha_1)^2 - \alpha_2] = 0$

(22) $E[R_t^e - \alpha_3] = 0$

(23) $E[(R_t^e - \alpha_3)^2 - \alpha_4] = 0$.

Again let $f(y_t, \alpha) = 0$ represent the above moment conditions, where y_t is now a $T\times 2$ matrix with data on gross equity returns and the IMRS, and α is a 4x1 vector of moments. Using (10'), the lower bound frontier for the case of one risky asset is given by:

$$\text{std}(m_v) = \sqrt{\frac{\left(1 - vER^e\right)^2}{\text{var}(R^e)}} .$$

The vertical distance measure is now written as:

(24) $$\gamma_2 \equiv h(\alpha) = \sqrt{\alpha_2} - \sqrt{\frac{(1-\alpha_1\alpha_3)^2}{\alpha_4}} .$$

The derivation of the asymptotic variance of both the α vector and the distance to the bound, γ_2, follows the procedure outlined in Section III.1. Under the null that the population vertical distance to the bound is zero, the Wald statistic for the one risky asset case is given by:

(25) $$Z_2 \equiv \sqrt{T}\left(\frac{\hat{\gamma}_2}{\hat{\sigma}_{\gamma_2}}\right) \xrightarrow{d} N(0,1) .$$

IV. Three Model Economies

Calculation of the distances γ_1 and γ_2 and the associated Wald statistics Z_1 and Z_2 requires information on the stochastic process for the intertemporal marginal rate of substitution, the return on the risky asset, and (for γ_1 and Z_1) the price of the riskless asset. In this section, we provide a description of how to obtain these stochastic processes for three model economies. Our three economies differ only in terms of preferences: the first has time-separable preferences (as in Lucas, 1978 and Mehra and Prescott, 1985), the second has state non-separable preferences (as in Epstein and Zin, 1989, 1991 and Weil, 1989), and the third has habit formation preferences (as in Sundaresan, 1989 and Constantinides, 1990). Though we use a common notation, none of the derivations here are new.

IV.1 Time-Separable Preferences

In this economy, there is a single tree that yields an exogenous stochastic flow of fruits, denoted by d_t at time t. The representative agent in this economy has preferences described by

$$(26) \qquad U_o = E_o \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}, \qquad \sigma > 0$$

where E_0 denotes conditional expectation given information at time 0, c_t denotes consumption at time t and $\beta \in (0,1)$ is the discount factor. (The $\sigma = 1$ case will be interpreted as logarithmic.) There is a competitive market for trading claims to the tree's fruits. The measure of agents and the measure of outstanding claims are each normalized to one, with the representative agent holding the single claim to the tree's fruits. With p_t denoting the price of one claim and s_t denoting the agent's shareholdings at time t, the agent's budget constraint is given by:

$$(27) \qquad c_t + p_t s_{t+1} \leq (p_t + d_t) s_t .$$

The agent's first-order conditions for choosing the optimal consumption and shareholding sequences are:

$$(28) \qquad c_t^{-\sigma} p_t = \beta E_t c_{t+1}^{-\sigma} (p_{t+1} + d_{t+1}), \quad t \geq 0.$$

In equilibrium all of the fruits are consumed each period and there is no other source of the consumption good, so $c_t = d_t$ for all t. The equilibrium prices are then determined as stationary functions of the state: $p_t = p(c_t)$. This method can be used to price assets with different payoff structures as well. For instance, the time-t price, p_t^f, of a one-period bond that pays one unit of consumption in period t+1 must satisfy

$$(29) \qquad c_t^{-\sigma} p_t^f = \beta E_t c_{t+1}^{-\sigma}, \qquad t \geq 0.$$

Following Mehra and Prescott (1985), we assume consumption growth to be a two-state Markov process with a symmetric transition matrix. We pick the three parameters of the Markov process for consumption growth, (the value of the "good" and "bad" states, denoted λ_1 and λ_2, and the probability of changing from state i to state j, denoted ϕ_{ij}, with symmetry implying that $\phi_{11} = \phi_{22}$), to match U.S. post-war quarterly per-capita real nondurables and services consumption. Using the methodology in Mehra-Prescott we can calculate the prices of the risky and riskless assets as functions of the state. (Details of the closed-form solutions for asset prices are relegated to the Appendix.) Thus, given a time series for consumption growth, we can obtain the time series for the price of riskless asset, equity return and the IMRS.

IV.2 State Non-Separable Preferences

Epstein and Zin (1989, 1991) and Weil (1989) generalized the time-separable preferences to allow for an independent parameterization of attitudes towards risk and intertemporal substitution. Following Weil (1989), we assume that the preferences are given by:

$$V_t = U[c_t, E_t V_{t+1}] \tag{30}$$

where

$$U[c, V] = \frac{\left\{(1-\beta)c^{1-\rho} + \beta[1+(1-\beta)(1-\sigma)V]^{\left(\frac{1-\rho}{1-\sigma}\right)}\right\}^{\left(\frac{1-\sigma}{1-\rho}\right)} - 1}{(1-\beta)(1-\sigma)}. \tag{31}$$

The elasticity of intertemporal substitution is $1/\rho$ and σ is the coefficient of relative risk aversion.

The economy in this model has the same structure as the one in Section IV.1. The Euler equation for the representative agent implies that for any asset with return $R_{k,t+1}$:

$$E_t\left\{\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}\right]^{\frac{1-\sigma}{1-\rho}}[R_{t+1}]^{\left(\frac{1-\sigma}{1-\rho}\right)-1} R_{k,t+1}\right\} = 1, \tag{32}$$

where R_{t+1} is the market return. To price trees (equity) we set $R_{k,t+1} = R_{t+1}$.

The IMRS for these preferences is:

$$\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}\right]^{\frac{1-\sigma}{1-\rho}}[R_{t+1}]^{\left(\frac{1-\sigma}{1-\rho}\right)-1}. \tag{33}$$

As in Section IV.1, given a time series for consumption growth, we can obtain the time series for the price of the riskless asset, equity return, and the IMRS. (See the Appendix for details.)

IV.3 Habit Formation Preferences

Sundaresan (1989), Constantinides (1990), and others model consumers who are habitual, in that levels of consumption in adjacent periods are complementary. That is, the preferences of consumers (in a discrete-time version of Constantinides, 1990) are given by:

$$U_o = E_o \sum_{t=0}^{\infty} \beta^t \frac{[(1+\delta(L))c_t]^{1-\sigma}}{1-\sigma}, \tag{34}$$

where $\delta(L)$ is a polynomial in the lag operator L. When the lag coefficients are all negative, the preferences exhibit habit-persistence. When the coefficients in $\delta(L)$ are zero, the preferences are time-separable. Here we work with the popular habit case with $\delta(L) = \delta L$.

The economy's structure again follows closely that of Section IV.1. The notation is a little more cumbersome since the marginal utility of consumption in period t must account for the fact that additional consumption in period t lowers utility in period t+1. The agent's first order condition for choosing optimal consumption and shareholding is:

$$[(c_t + \delta c_{t-1})^{-\sigma} + \beta\delta(c_{t+1} + \delta c_t)^{-\sigma}]p_t = \beta E_t[(c_{t+1} + \delta c_t)^{-\sigma} + \beta\delta(c_{t+2} + \delta c_{t+1})^{-\sigma}](p_{t+1} + d_{t+1}) . \tag{35}$$

Again $c_t = d_t$ in equilibrium. The representative agent's IMRS is given by:

$$m_{t+1} = \beta \frac{(c_{t+1} + \delta c_t)^{-\sigma} + \beta\delta E_{t+1}(c_{t+2} + \delta c_{t+1})^{-\sigma}}{(c_t + \delta c_{t-1})^{-\sigma} + \beta\delta E_t(c_{t+1} + \delta c_t)^{-\sigma}} . \tag{36}$$

Time series for the price of the riskless asset, equity return, and the IMRS again follow from the time series for consumption growth. (See the Appendix for the details on the calculation of asset prices.)

V. Monte Carlo Results for Time-separable Preferences

In the previous literature (Gregory and Smith, 1992, and Burnside, 1994, for example), the usefulness of the HJ bound as an evaluation device has been judged by simulating the asset pricing model with time-separable preferences many times and examining the frequency of 'rejections' when the model is in fact true. The simplest evaluation metric is the fraction of times the model IMRS violates the HJ bound. A metric that takes into account sampling variability involves rejection frequencies based on asymptotic critical values for Z_1 and Z_2 that emerge from the distribution theory in Section III. In the remainder of the paper we focus on the statistical measures of the distance to the bound. (If one uses the finite-sample distributions of the test statistics, then the frequency of rejections is pinned down by choice of the finite-sample critical values.)

In this section, we begin by presenting the procedure for our Monte Carlo simulations of the model with time-separable preferences. We then present some evidence on the difference between the small-sample and asymptotic distributions of the test statistics. We then document the dependence of the Z_1 and Z_2 distributions on nuisance parameters, calculate critical values that deliver rejection rates no greater than desired regardless of nuisance parameters. In subsequent sections we show that the critical values are valid for state non-separable and habit formation preferences as well.

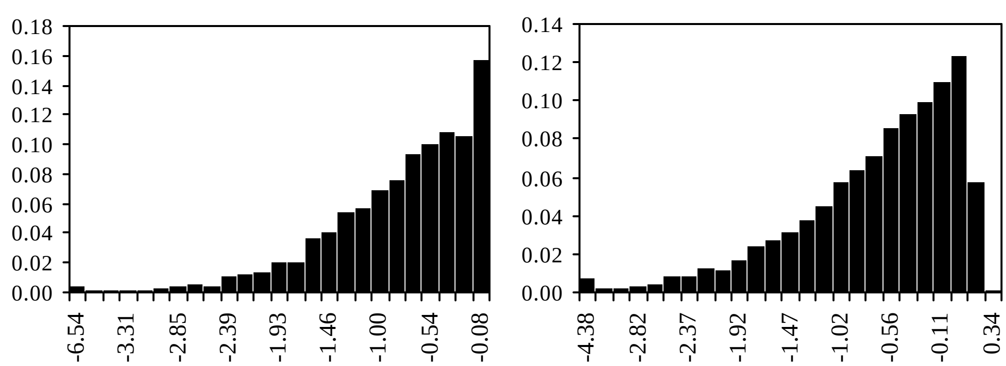
Our investigation is based on the following simulation procedure:

1. Draw a consumption time series of length 200, say, for the two-state Markov process described in Section IV.1;
2. Calculate the IMRS time series implied by the model with time-separable preferences;
3. Price the two assets, equity and a riskless bond, and calculate the time series for asset returns;
4. Calculate the HJ bound and the standard deviation of the model's IMRS;
5. Calculate the distance between the model and the HJ bound and the value of the Wald statistic;
6. Repeat steps 1-5 for 1000 different draws of the consumption time series.

Preliminary evidence on the small-sample distribution of the Z_1 statistic is presented in Figure 1, which plots estimated density functions for 4 different values of σ when there is 1 riskless asset and 1 risky asset.[5] Note that for small σ, most of the distribution is to the left of zero (meaning that the bound is nearly always violated in simulations from the true model), so we would reject the true model frequently if we were to use a 'naïve' metric that did not account for sampling variability. In all cases it is clear that the distribution is not Standard Normal, and hence we should not use the asymptotic critical values. Furthermore, the distribution changes as risk aversion changes, an issue we will investigate more fully in subsequent subsections.

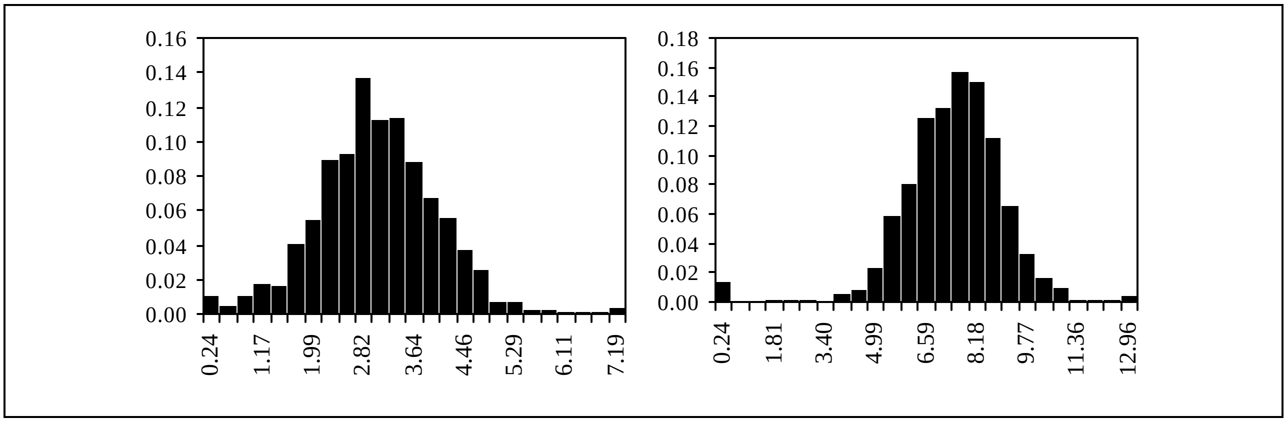
Figure 1: PDFs of the Z_1 Statistic for 1 Riskless Asset and 1 Risky Asset

a: σ = 0.005 b: σ = 2



c: σ = 50 d: σ = 100

[5] Burnside (1994) also examined the small-sample properties of the Wald statistic using time-separable preferences with one risky asset and one riskless asset and computed the critical values. The version of the HJ bound he used, however, is the one for the case with *no* riskless asset. In population, his bound will always be greater than ours, and thus provides a more restrictive test than ours.



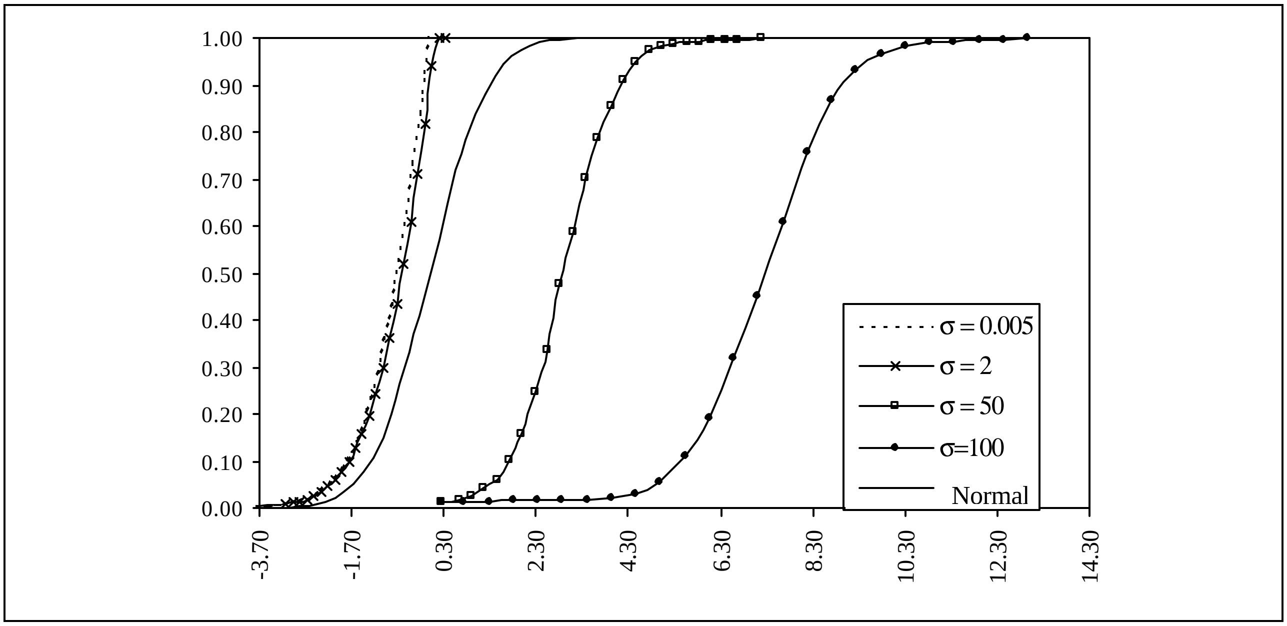
V.1 Dependence of finite-sample distribution of Z_1 on s

In Figure 2 below, we illustrate the dependence on σ of the cumulative distribution functions for the Z_1 statistic. Analogous calculations with fixed σ and different values of β revealed that changing the rate of time preference has a negligible effect on the Z_1 distribution. There are three features worth noting in the figure. First, for small σ, the distribution is clearly not Normal, but as σ rises, the distribution approaches Normality.[6] The second feature is the rightward movement as σ increases. This dependence of the distribution on σ means that the Z_1 test is *non-similar*. The relevant critical value of interest is the one corresponding to the lowest σ because the bound is satisfied in population for all possible σ, and we are interested in a lower one-tail test.[7] Since the distributions move rightward as σ increases, using the critical values from the smallest σ distribution (i.e., the leftmost one) ensures that the size of the type I error will be no larger than designed (and could be substantially smaller). Finally, the third feature of the figure is that for small σ the CDFs lie to the left of the Standard Normal distribution, so the asymptotic critical value would be larger than that (to the right of) of the small sample, leading to overrejection.

Figure 2: CDFs of the Z_1 Statistic

[6] A formal test indicates that the skewness and kurtosis are statistically different from those of the Normal distribution. This is true for all distributions that we present subsequently.

[7] When $\sigma = 0$, the IMRS is equal to β in every state, so its variance is zero. For an interior solution, $p^f = \beta$ and $ER^e = 1/p^f$ so from (16), var(m*) = 0 also. Thus the bound is just attained when $\sigma = 0$.



To calculate the critical value appropriate for a sample of length 200, we utilize a sample of 10000 time series for the case $\sigma = 0.005$. (The boundary value of $\sigma = 0$ associated with the risk-neutral case is problematic because of indeterminacies that arise in pricing assets with linear indifference curves. Reducing σ further to 0.0001, a factor of 50, caused negligible changes in critical values.) The 5% critical value is -2.1258, clearly quite different from the Standard Normal (asymptotic) critical value of -1.65. If one uses the asymptotic critical value, the rejection rate in this sample would be more than 10%. Moreover, using the extreme-σ 5% finite-sample critical value in cases in which σ is actually larger yields type I error rates much smaller than 5%: for instance, when $\sigma = 50$, the rejection rate is 0%.

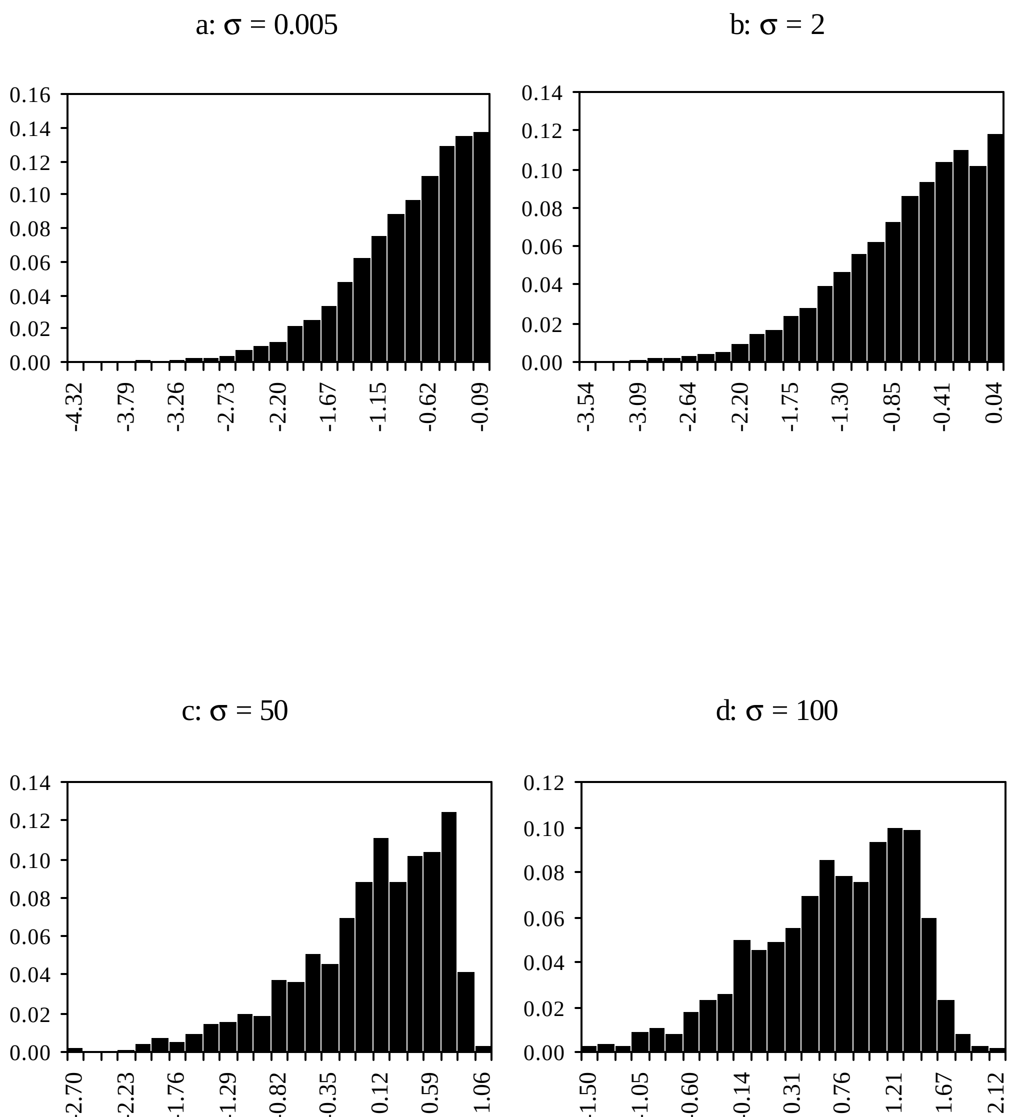
The calculations in this subsection involve the rather artificial situation in which there is an identifiable riskless asset. In the next sub-section, we repeat these calculations for an economy in which there is no riskless asset and investigate the properties of the Wald statistic Z_2.

V.2 Dependence of finite-sample distribution of Z_2 on s

In the case where there is no riskless asset, restriction (3) does not apply. The returns version of the HJ bound frontier is now given by (10') in Section II.2 and the Z_2 statistic is given by (25) in Section

III.2. We consider the case where there is only one risky asset.[8] Figure 3 plots the density functions of the Z_2 statistic for the four values of σ.

Figure 3: PDFs of Z_2 Statistic for 1 Risky Asset



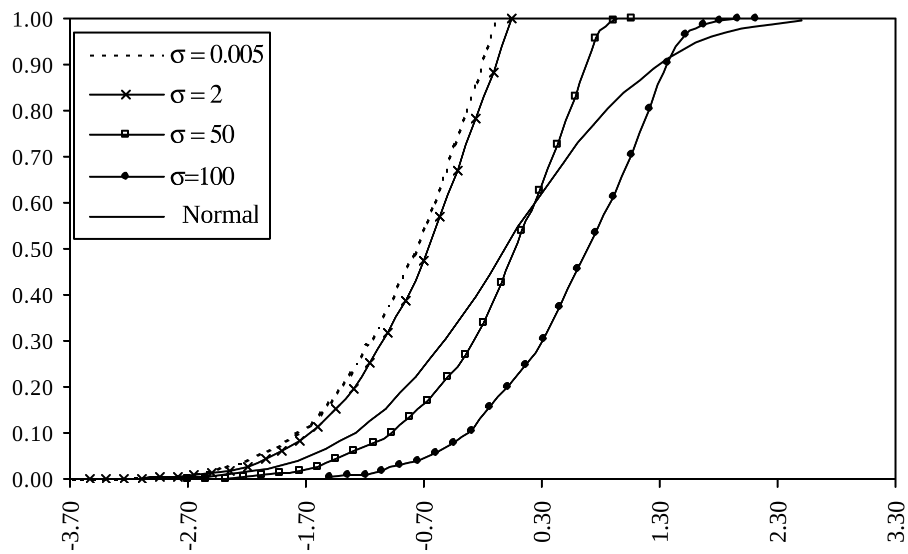
The PDFs do not appear to be dramatically different from the case with a riskless asset. The PDFs are clearly non-Normal. Further, as σ rises, the distributions migrate rightward, though this

[8] When generating a model economy, we can easily price one risky asset (one "tree") by setting dividends from the asset equal to consumption. If there are multiple risky assets, we will have to take a stand on how the aggregate dividends are distributed across the risky assets.

movement is not as pronounced as in the riskless-asset case. Figure 4 plots the CDFs associated with the above four density functions.

The finite-sample 5% critical value for Z_2 is -1.9867. As in the case of Z_1, this is very different from the asymptotic critical value. In fact, if one were to use the asymptotic critical value, the rejection rate would be 10.5%. Moreover, using the extreme-σ 5% finite-sample critical value in cases in which σ is actually larger yields type I error rates much smaller than 5%: for instance, when $\sigma = 2$, the rejection rate is 4.5% and when $\sigma = 50$, the rejection rate is 0.5%.

Figure 4: CDFs of the Z_2 Statistic



VI. Non-Separable Preferences

For time-separable preferences, we documented in the previous section that as σ decreases, the CDFs of the Wald statistics shift to the left. We argued that the leftmost distribution is the most appropriate one to use in order to calculate small-sample critical values since such critical values produce maximal size across (unknown) values of σ. As is well known, for CRRA preferences the parameter σ describes two attributes of agents' preferences: risk aversion and intertemporal substitution. State non-separable preferences relax this assumption and allow risk aversion and intertemporal substitution to be altered independently. Similarly, habit formation preferences also sever the tight link between intertemporal substitution and risk aversion that characterizes the time-separable case; for example, such preferences allow for elasticities of intertemporal substitution that vary with the levels of current and past consumption. Both forms of non-separabilities are popular in preference-based asset-pricing models. In this section, we investigate the small-sample properties of Z_1 and Z_2 statistics for economies with non-separable preferences.

VI.1 State Non-Separable Preferences

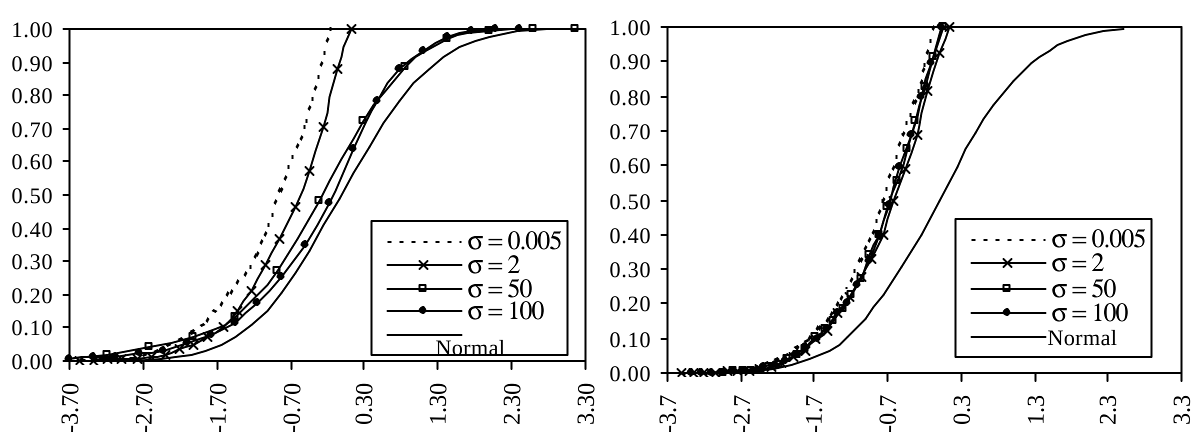
By employing a Monte Carlo study of the effect of varying risk aversion and intertemporal substitution separately, we can determine which effect, low risk aversion or high intertemporal substitution, causes the leftward migration of the CDFs of Z_1 and Z_2. Equations 30-31 describe the Epstein-Zin preferences that allows us parameterize risk aversion and intertemporal substitution separately.

Figure 5 displays the CDFs for the Wald statistics Z_1 and Z_2 for ρ (the inverse of the elasticity of intertemporal substitution) fixed at 0.005, and σ (risk aversion) varying from 0.005 to 100. We see in panel a that as we raise risk aversion in the 1 risky, 1 riskless asset case, the CDFs move rightward, yet they do not appear to vary much with σ. In panel b, the no riskless asset case, the distributions appear the same regardless of σ.

Figure 5: CDFs of Z_1 and Z_2 for Fixed Intertemporal Substitution

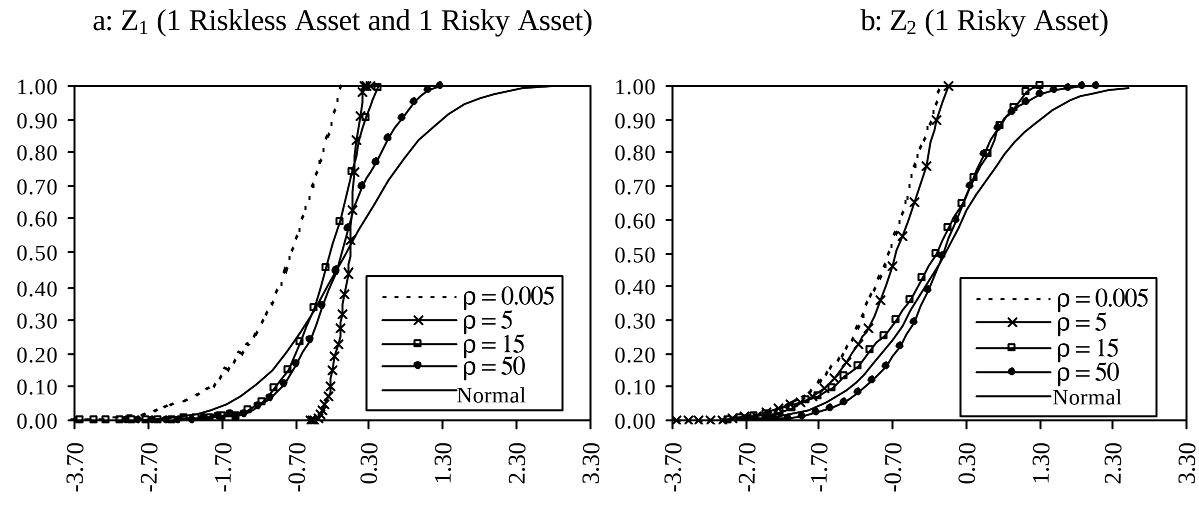
a: Z_1 (1 Riskless Asset and 1 Risky Asset)

b: Z_2 (1 Risky Asset)



Next we fix σ and vary ρ between 0.005 and 100. In panel a of Figure 6, the one risky and one riskless asset case, the movement in the CDFs is non-monotonic. A similar phenomenon characterizes the no riskless asset case depicted in panel b.

Figure 6: CDFs of Z_1 and Z_2 for Fixed Risk Aversion



In both Figure 5 and Figure 6, the small-sample distributions are non-Normal and the asymptotic critical values are inappropriate for evaluating the asset pricing model. In contrast, with high intertemporal substitution ($1/\rho = 200$), we get distributions very much like the time-separable case, regardless of risk aversion. With low risk aversion ($\sigma = 0.005$), there is a lot of movement in the CDFs as intertemporal substitution is decreased. That is, the small-sample distribution of the Wald statistic is influenced significantly by the degree of intertemporal substitution, and not by attitudes toward risk. Further, we can conclude that the distribution to use in order to calculate the maximal-size critical value is the one corresponding to high intertemporal substitution and low risk aversion. Thus, we can simply use the small-sample critical values associated with time-separable preferences.

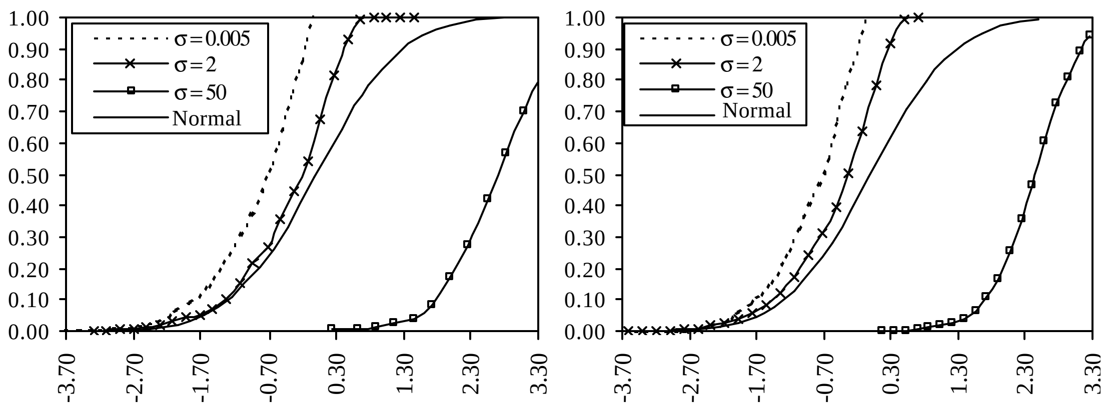
VI.2 Habit Formation Preferences

Here we document the relative importance of the curvature parameter and the non-separability parameter in determining the small-sample properties of the test statistic. Further, we argue that the maximal-size critical values for time-separable preferences are appropriate for habit formation preferences as well. See equation 34 for more on habit formation preferences.

First we fix δ at -0.5 and vary σ between 0.005 and 50. Recall that with $\delta = 0$ the preferences are time-separable; $\delta = -0.5$ is far enough from the time-separable case to illustrate the effect of habit formation. As we shall see, the exact value of δ will not affect our conclusions. In panel a of Figure 7, the one risky asset and one riskless asset case, the distribution of Z_1 moves rightward with σ, as in the case of time-separable preferences. For the no riskless asset case, panel b of Figure 7, the distributions of Z_2, shift rightward as σ rises. In both figures, the effect of the curvature parameter σ is qualitatively the same as in the time-separable case, the only difference being that the rightward shifts in the distribution are noticeable for smaller values of the curvature parameter.

Figure 7: CDFs of Z_1 and Z_2 for Fixed Habit Persistence

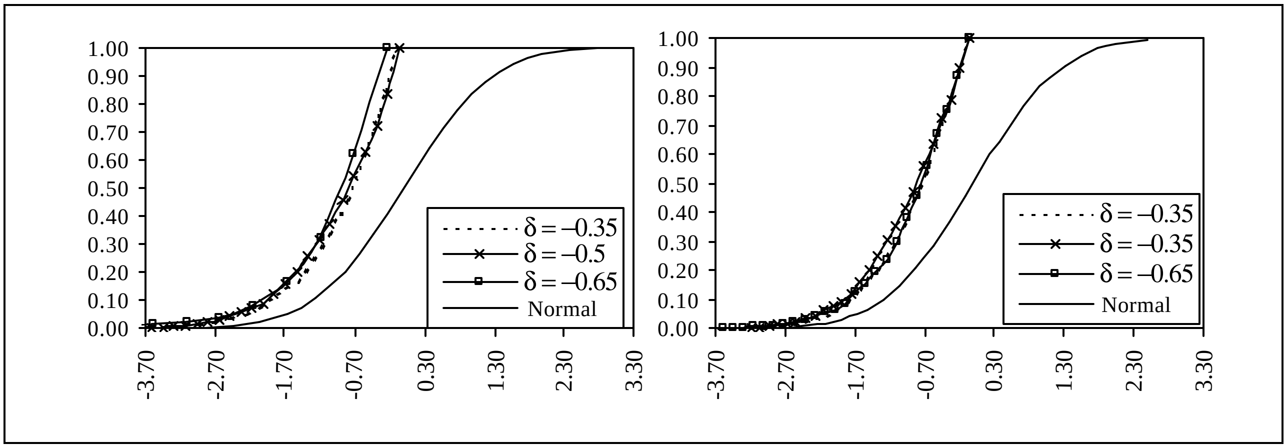
a: Z_1 (1 Riskless Asset and 1 Risky Asset) b: Z_2 (1 Risky Asset)



In Figures 8a and b we show the effect of changes in δ when σ is very small (0.005). In this case, the effect of changing δ is not visible. That is, the CDFs look very much like those in the time-separable case. Further, in both Figure 7 and Figure 8, the small-sample distributions are non-Normal, so the rejection rates based on asymptotic critical values are inaccurate.

Figure 8: CDFs of Z_1 and Z_2 for Fixed σ

a: Z_1 (1 Riskless Asset and 1 Risky Asset) b: Z_2 (1 Risky Asset)



What we take from all these calculations is that the proper, maximal-size, small-sample critical values for the habit model should be calculated with small σ, regardless of the value for δ. This result is important in light of the large number of functional forms used to model time-non-separabilities. All the functional forms are of the general class: $U_o = E_o \sum_{t=0}^{\infty} \beta^t \frac{[f(c_t, c_{t-1}, \ldots)]^{1-\sigma}}{1-\sigma}$, where the function $f(\cdot)$ relates current consumption to past consumption in some manner. Note that this general form contains both 'internal' and 'external' habit.[9] Our results suggest that for any of these forms, one need only use the critical values from the time-separable model.

VII. Maximal-Size Critical Values

We have argued that the critical values generated from a Monte Carlo study with time-separable preferences with near-linear period utility are appropriate for evaluating time separable, time non-separable and state non-separable preference-based asset pricing models. The result was obtained in a simple two-state Markov economy, which is convenient when working with time-non-separable preferences. Here we calculate small-sample critical values for the Wald statistics for a consumption process that is designed to match more characteristics of U.S consumption data than is possible in the two-state framework.

The model economy is the same one described in Section IV.1. However, we allow for many more possible states. We retain the discrete Markov process so that we can obtain closed-form solutions for asset prices, but allow for more states to get a better approximation of the underlying continuous

[9] The function $f(\cdot)$ could simply be a longer lag polynomial, but of the same form used here. A few other possibilities include Heaton (1995), who places restrictions on the infinite-order lag polynomial parameters, Abel (1990, 1998) and Campbell and Cochrane (1998) who develop non-linear functions for $f(\cdot)$.

process. Tauchen (1986a) describes a procedure for choosing values for the states and state transition probabilities. His procedure has the property that as the number of states rises, the discrete process converges to the continuous process. Tauchen (1986b) uses the procedure to simulate a model economy (identical to ours) and to study the small-sample properties of GMM estimators of utility function parameters. He finds that the approximation procedure works well.

There are two issues when approximating the U.S. consumption growth process. First, what is a 'good' time series representation for, say, quarterly consumption growth? We have estimated AR(1) through AR(5) processes, and AR(3) seems appropriate. In the AR(4) and AR(5) estimates the coefficients on lags greater than 3 are small and not statistically different from zero. Further, when we repeatedly simulate the AR(3) process we find that the small-sample distribution of the first 3 moments and first 4 autocorrelations center on the sample estimates of U.S. consumption growth for these moments. When we simulate the AR(2) process this is not the case.

The second issue is the number of discrete states needed to approximate the continuous AR(3) process. To approximate the process we begin by writing the consumption AR(3) process as a VAR(1). We then allow N = 9 possible values for consumption growth, meaning that there are $N^3 = 729$ states in the VAR(1).[10] We simulate the Markov economy and compare the distribution of the first 4 moments and 6 autocorrelations with the distribution we obtained by simulating the continuous AR(3) process. In all cases the distributions were very close.

The procedure for calculating the small-sample distribution of the Wald statistics is the same as in Section V. We simulate the economy with time separable preferences and $\sigma = 0.005$, calculate asset prices and returns in each of the 729 states, and calculate the Wald statistics Z_1 and Z_2. We repeat this for time series of various lengths. In Table 1, we provide critical values for consumption growth parameterized to post-World War II quarterly data. There are two striking features of the table. The first is that there does not appear to be much difference between the case with a riskless asset and the case without the riskless asset. This was not true when we used a two-state Markov process for consumption growth. The second feature is that even for sample sizes that are large relative to typical empirical applications, the finite-sample critical value is different from the asymptotic critical value.

[10] Tauchen (1986a) provides a method for choosing the grid values and probabilities for a general VAR(1) process. Here, the VAR(1) actually represents a univariate AR(3) with an expanded state space (the extra 2 elements being represented by the immediate past two values of the process); this means, for example, that consumption growth at time t-1 is known at time t. This feature requires us to employ suitable modifications of the transition probabilities presented in Tauchen (1986a).

We also calculated critical values using annual data from 1889 to 1992. Surprisingly, there are few differences between the critical values for the two cases. Given the dramatic changes in the consumption growth process between the two periods (see Golob, 1992 or Otrok, Ravikumar, and Whiteman, 2000), this result is somewhat surprising. In any case, it is comforting that the critical values in Table 1 can be used without much concern for the properties of consumption growth in a particular subsample.

Table 1: Small-sample Critical Values: Z_1 and Z_2

Consumption Growth Parameterized to Quarterly U.S. Data, 1947:1-1997:4

	Z_1 (One Riskless and One Risky Asset)				Z_2 (One Risky Asset)			
Quantile	**T=100**	**T=200**	**T=300**	**T=500**	**T=100**	**T=200**	**T=300**	**T=500**
1.0%	-2.892	-2.690	-2.652	-2.676	-2.904	-2.690	-2.641	-2.674
2.5%	-2.489	-2.311	-2.265	-2.344	-2.490	-2.305	-2.261	-2.333
5.0%	-2.145	-2.040	-2.002	-2.014	-2.134	-2.035	-2.004	-2.010
10.0%	-1.776	-1.692	-1.665	-1.683	-1.772	-1.691	-1.666	-1.682
25.0%	-1.222	-1.165	-1.166	-1.161	-1.221	-1.164	-1.165	-1.162
50.0%	-0.724	-0.691	-0.685	-0.671	-0.723	-0.690	-0.684	-0.671
75.0%	-0.336	-0.330	-0.326	-0.324	-0.336	-0.330	-0.327	-0.323
90.0%	-0.134	-0.136	-0.131	-0.125	-0.134	-0.136	-0.131	-0.125
95.0%	-0.065	-0.066	-0.066	-0.066	-0.065	-0.066	-0.066	-0.066
97.5%	-0.030	-0.034	-0.033	-0.031	-0.030	-0.034	-0.033	-0.031
99.0%	-0.013	-0.012	-0.012	-0.013	-0.013	-0.012	-0.012	-0.013

VIII. An Application to U.S. Data

In this section we use the small-sample critical values presented in the previous section to 'test' the three asset pricing models described in Section IV using U.S. consumption and return data from 1947:1-1997:4. For the time separable model with logarithmic preferences the Z_1 statistic is -13.131, suggesting a rejection of the model. Significantly larger values for the curvature parameter do not lead to acceptance of the model. For example, a curvature parameter of 50 leads to a Z_1 statistic of -12.861, leading to a rejection of the model again.

The Epstein-Zin model fares better than the time-separable model for some parameter values. With $\rho = 1.3$ (the inverse of intertemporal substitution) and $\sigma = 5$ (risk aversion), the Z_1 statistic is -12.309, indicating a model rejection. However, holding intertemporal substitution constant and increasing risk aversion to 14, the Z_1 statistic is -0.1105, so we would fail to reject this model. That is, for plausible values of intertemporal substitution and risk aversion, and after accounting for sampling error, we would conclude that the Epstein-Zin model is not inconsistent with historical return data. Of course, the HJ bound is only a necessary condition and we would want to study the model more carefully before accepting it.

The results for the habit model are mixed. With $\delta = -0.712$ and $\sigma = 7.12$, values that are consistent with most of the literature on habit formation, the value of the Z_1 statistic is -1.4184, so we would not reject this model at conventional significance levels. For a more moderate amount of habit, say $\delta = -0.2$, and $\sigma = 10$, the Z_1 statistic is -14.087, which would call for a rejection of the model. These mixed results have a number of implications. First, because the model is soundly rejected with the latter set of parameters, we would not want to go through the effort to more carefully analyze that version of the model. Second, since the habit model with the first set of parameters is not rejected, we would want to subject this model to closer scrutiny. Again, this is because the HJ bound is only a necessary condition. In Otrok, Ravikumar and Whiteman (2000) we develop a spectral approach to evaluate and understand the habit model. We conclude that the model with these parameters has a number of troublesome predictions.

IX. Conclusion

We have investigated the usefulness of the Hansen-Jagannathan (1991) bound for assessing the empirical plausibility of asset pricing models. For completeness, we provided a derivation of the bound, the asymptotic distribution theory for a distance measure based on the bound, and a description of three sources of IMRSs—time separable preferences, state non-separable preferences, and habit-formation preferences. We studied how the bound can be used to assess the three asset pricing models with data samples of the size typically encountered in applied work.

For the model with time-separable preferences, we showed that the Hansen-Jagannathan distance measure test is *non-similar* in finite samples. Under the null hypothesis that the model is true, the distributions of the test statistics depend in a nontrivial way on the curvature parameter of the utility function. We then showed that inference based on critical values taken from the almost linear utility (risk neutral) case yields type I errors no larger than intended across any positive values for the curvature parameter; i.e., our critical value calculations are designed to guarantee 'small' type I errors. We also documented that our critical values for the risk-neutral case are valid for the two non-separable preferences as well.

Next, we provided finite-sample critical values for the distance measure using a high-dimension Markov process for consumption growth. We conclude that when these are used, the HJ bound is in fact a useful device for evaluating asset pricing models. Finally, we investigated the three asset pricing models using U.S. annual data and our small-sample critical values. We found evidence against the time-separable model, and mixed evidence for the time and the state non-separable preferences.

References

Abel, Andrew B., (1990), "Asset Pricing Under Habit Formation and Catching Up With The Joneses," *American Economic Review* 80: 38-42.

Abel, Andrew B., (1998), "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics*, February 1999, v. 43, No. 1, pp. 3-33.

Burnside, Craig, (1994), "Hansen-Jagannathan Bounds as Classical Tests of Asset Pricing Models," *Journal of Business and Economic Statistics* 12: 57-79.

Campbell, John Y., and John H. Cochrane, (1998), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, April 1999, v. 107, No. 2, pp. 205-51.

Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark, (1994), "Testing Volatility Restrictions on Intertemporal Marginal Rates of Substitution Implied by Euler Equations and Asset Returns," *Journal of Finance* 49: 123-152.

Cochrane, John H., Lars P. Hansen, (1992), "Asset Pricing Explorations for Macroeconomics," in Olivier J. Blanchard and Stanley Fischer, eds. *NBER Macroeconomics Annual 1992* 115-163.

Constantinides, George C., (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy* 98: 519-43.

Epstein, Larry G., and Stanley E. Zin, (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57: 937-968.

Epstein, Larry G., and Stanley E. Zin, (1991), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy* 99: 263-286.

Golob, John, (1992), "A Regime Shift in Measured Per Capita Consumption, with Implications for Asset Prices and Returns," Federal Reserve Bank of Kansas City, research working paper RWP 92-06.

Gregory, Allan W., and Gregor W. Smith, (1992), "Sampling Variability in Hansen-Jagannathan Bounds," *Economics Letters* 38: 263-267.

Hansen, Lars P., (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50:1029:1054.

Hansen, Lars P., John Heaton and Erzo G.J. Luttmer, (1995), "Econometric Evaluation of Asset Pricing Models," *The Review of Financial Studies* 8: 237-274.

Hansen, Lars P., and Ravi Jagannathan, (1991), "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy* 99: 225-262.

Hansen, Lars P., and Ravi Jagannathan, (1997), "Assessing Specification Errors in Stochastic Discount Factor Models," *The Journal of Finance* 52: 557-590.

Heaton, John, (1995), "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," *Econometrica* 63:3, 681-717.

Lucas, Robert E. Jr., (1978), "Asset Prices in an Exchange Economy," *Econometrica* 46: 1429-1445.

Mehra, Rajnish and Edward C. Prescott, (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15: 145-161.

Otrok, Christopher, B. Ravikumar and Charles H. Whiteman, (2000), "Habit Formation: A Resolution of the Equity Premium Puzzle?," working paper, University of Iowa.

Sundaresan, Suresh M., (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies* 2: 73-89.

Tauchen, George, (1986a), "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters* 20: 177-181.

Tauchen, George, (1986b), "Statistical Properties of Generalized Method of Moments Estimators of Structured Parameters Using Financial Market Data," *Journal of Business and Economic Statistics*, 4: 397-416.

Weil, Philippe, (1989), The Equity Premium Puzzle and The Risk-Free Rate Puzzle," *Journal of Monetary Economics* 24: 401-421.

Appendix: Asset Prices

A.1 Time Separable Preferences. Mehra and Prescott (1985) obtain closed-form solutions for asset prices as follows. From (25) the price of the risky asset is:

$$\text{(A1)} \qquad p(c,i) = \beta \sum_{j=1}^{2} \phi_{ij} \frac{(\lambda_j c)^{-\sigma}}{c^{-\sigma}} [p(\lambda_j c, j) + \lambda_j c].$$

To solve for the unknown prices, Mehra and Prescott use the fact that prices are homogenous of degree one in consumption and write prices as: $p(c,i) = w_i c$. Substituting in (A1) yields a set of two equations with two unknowns:

$$\text{(A2)} \qquad w_i = \beta \sum_{j=1}^{2} \phi_{ij} \lambda_j^{1-\sigma} [w_j + 1], \; i = 1, 2.$$

The formula for the price of the riskless asset in state i is:

$$\text{(A3)} \qquad p_i^f = \beta \sum_{j=1}^{2} \phi_{ij} \frac{(\lambda_j c)^{-\sigma}}{c^{-\sigma}} = \beta \sum_{j=1}^{2} \phi_{ij} \lambda_j^{-\sigma}.$$

The gross return on equity is $R_{t+1} = \dfrac{p_{t+1} + c_{t+1}}{p_t}$; the return on the riskless asset is $R_{t+1}^f = \dfrac{1}{p_t^f}$.

A.2 State Non-Separable Preferences. The w's now solve:

$$\text{(A4)} \qquad w_i = \beta \left\{ \sum_{j=1}^{2} \phi_{ij} \lambda_j^{1-\sigma} (w_j + 1)^{\frac{1-\sigma}{1-\rho}} \right\}^{\frac{1-\rho}{1-\sigma}}.$$

The price of the riskless asset in state i is:

$$\text{(A5)} \qquad p_i^f = \beta^{\frac{1-\sigma}{1-\rho}} \left\{ \sum_{j=1}^{2} \phi_{ij} \lambda_j^{-\sigma} \left(\frac{w_j + 1}{w_i}\right)^{\frac{\rho-\sigma}{1-\rho}} \right\}.$$

The unknown w's are found using a nonlinear equation solver in Gauss.

A.3 Habit Formation Preferences. Following the same procedure as in Section IV.1, the price of the risky asset in state 1 can be written as:

$$
\text{(A6)}\quad
\begin{aligned}
p(c,1) = \beta\phi_{11}\phi_{11}&\left(\frac{(\lambda_1 c+\delta c)^{-\sigma}+\beta\delta(\lambda_1\lambda_1 c+\delta\lambda_1 c)^{-\sigma}}{(c+\delta\frac{c}{\lambda_1})^{-\sigma}+\beta\delta(\lambda_1 c+\delta c)^{-\sigma}}\right)\left(p(\lambda_1 c,1)+c\lambda_1\right)+\\
\beta\phi_{11}\phi_{12}&\left(\frac{(\lambda_1 c+\delta c)^{-\sigma}+\beta\delta(\lambda_2\lambda_1 c+\delta\lambda_1 c)^{-\sigma}}{(c+\delta\frac{c}{\lambda_1})^{-\sigma}+\beta\delta(\lambda_1 c+\delta c)^{-\sigma}}\right)\left(p(\lambda_1 c,1)+c\lambda_1\right)+\\
\beta\phi_{12}\phi_{21}&\left(\frac{(\lambda_2 c+\delta c)^{-\sigma}+\beta\delta(\lambda_1\lambda_2 c+\delta\lambda_2 c)^{-\sigma}}{(c+\delta\frac{c}{\lambda_1})^{-\sigma}+\beta\delta(\lambda_2 c+\delta c)^{-\sigma}}\right)\left(p(\lambda_2 c,2)+c\lambda_2\right)+\\
\beta\phi_{12}\phi_{22}&\left(\frac{(\lambda_2 c+\delta c)^{-\sigma}+\beta\delta(\lambda_2\lambda_2 c+\delta\lambda_2 c)^{-\sigma}}{(c+\delta\frac{c}{\lambda_1})^{-\sigma}+\beta\delta(\lambda_2 c+\delta c)^{-\sigma}}\right)\left(p(\lambda_2 c,2)+c\lambda_2\right).
\end{aligned}
$$

There is an analagous price equation for state 2. Substituting $p(c,i) = w_i c$ into (A6) results in two equations and two unknowns, which can be solved for the prices in each state, as in the case of time-separable preferences. The price of the riskless assset in the habit case is calculated by substituting a 1 for $(p(\lambda_i c,i)+c\lambda_i)$ in equation (A6).